Genetic Circuitry of Survival Motor Neuron, the Gene Underlying Spinal Muscular Atrophy

The genetic circuitry of Survival Motor Neuron, the gene underlying Spinal Muscular Atrophy

Anindya Sen[1]*, Douglas N. Dimlich[1]*, K. G. Guruharsha[1]*, Mark W. Kankel[1]*, Kazuya Hori[1], Takakazu Yokokura[1,2], Sophie Brachat[3,4], Delwood Richardson[3], Joseph Loureiro[3], Rajeev Sivasankaran[3], Daniel Curtis[3], Lance S. Davidow[5], Lee L. Rubin[5], Anne C. Hart[6], David Van Vactor[1], and Spyros Artavanis-Tsakonas[1]. * Equal contribution

1.Department of Cell Biology, Harvard Medical School, Boston, MA 02115, USA. 2.Current affiliation: Okinawa Science and Technology Graduate University, 1919-1 Tancha, Onna-son, Okinawa 904-0495, Japan 3.Developmental and Molecular Pathways, Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. 4.Current affiliation: Musculoskeletal Diseases, Novartis Institutes for Biomedical Research, Novartis Campus, CH-4002 Basel, Switzerland 5.Department of Stem Cell and Regenerative Biology, Harvard Medical School, Boston, MA 02115, USA. 6.Department of Neuroscience, Brown University, 185 Meeting Street Box GL-N, Providence, RI 02912, USA.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

The clinical severity of the neurodegenerative disorder Spinal Muscular Atrophy (SMA) is dependent on the levels of functional Survival Motor Neuron (SMN) protein. Consequently, current strategies for developing treatments for SMA generally focus on augmenting SMN levels. To identify additional potential therapeutic avenues and achieve a greater understanding of SMN, we applied in vivo, in vitro, and in silico approaches to identify genetic and biochemical interactors of the Drosophila SMN homolog. We identified more than three hundred candidate genes that alter an Smn-dependent phenotype in vivo. Integrating the results from our genetic screens, large-scale protein interaction studies and bioinformatics analysis, we define a unique interactome for SMN which provides a knowledge base for a better understanding of SMA.

Genetic screen | Interactome | Proteomics | Spinal Muscular Atrophy | Survival Motor Neuron

INTRODUCTION

Spinal Muscular Atrophy (SMA), the leading genetic cause of infant mortality, results from the partial loss of Survival Motor Neuron (SMN) gene activity (1). Numerous studies indicate that SMN functions as a central component of a complex which is responsible for the assembly of spliceosomal small nuclear ribonucleoproteins (snRNPs) [reviewed in (2)]. SMN is also reported to play additional roles, including mRNA trafficking in the axon (3). In humans, SMN is encoded by two nearly identical genes, *SMN1* and *SMN2*, which are located on chromosome 5 (4). *SMN2* differs from *SMN1* in that only 10% of *SMN2* transcripts produce functional SMN due to a single nucleotide polymorphism that results in inefficient splicing of exon 7 and translation of a truncated, unstable SMN protein (1, 5, 6). The clinical severity of SMA correlates with *SMN2* copy number, which varies between individuals (7). As the small amount of functional SMN2 protein produced by each copy is capable of partially compensating for the loss of the *SMN1* gene function, higher copy numbers of *SMN2* typically result in milder forms of SMA. Therefore, genetic modifiers capable of increasing the abundance and/or specific activity of SMN hold promise as therapeutic targets.

The *Drosophila* genome harbors a single, highly conserved ortholog of SMN1/2, the *Survival motor neuron* (*Smn*) gene. SMN is essential for cell viability in vertebrates and *Drosophila* (8, 9). In *Drosophila*, zygotic loss of *Smn* function results in recessive larval lethality (not embryonic as might be expected) due to the rescue of early development by maternal contribution of *Smn*. The larval phenotype includes neuromuscular junction (NMJ) abnormalities that are reminiscent of those associated with the human disease, rendering this invertebrate organism an excellent system to model SMN biology (10-12). We previously described a genetic screen for modifiers of the lethal phenotype resulting from a complete loss of function *Smn* allele (13). This screen, though it probed half of the *Drosophila* genome, identified only a relatively small number of genes that affected the NMJ phenotype associated with *Smn* loss of function (13). In particular, it did not identify genes involved in snRNP biogenesis, the molecular functionality that is most clearly associated with SMN.

As the human disease state results from partial loss of SMN function, we reasoned that a screening paradigm using a hypomorphic *Smn* background, (as opposed to a background that completely eliminates SMN function) would more closely resemble the genetic condition in SMA. Such a screen would therefore enhance our ability to detect novel elements of the *Smn* genetic network, and, consequently, add significantly to our efforts to both dissect the *Smn* genetic circuitry as well as identify potential clinically relevant targets with novel mode of action.

This complementary screen proved to be more sensitive than our previous screen and led to the identification of over 300 genetic interactors. Taking advantage of the recently established *Drosophila* Protein Interaction Map (DPiM) (14), we related the newly identified genetic interactors to the SMN protein interactome, producing an integrated *Drosophila* SMN biological network. Finally, the *Drosophila* SMN network was evaluated for its relevance to human biology by mapping *Drosophila* SMN network genes to their human homologs, and analyzing the human network using computational biology tools. The projection of the *Drosophila* SMN network derived from this study onto the human network derived from prior knowledge provides a rational basis for novel SMN functional hypotheses and network

Significance

Spinal Muscular Atrophy (SMA), the leading genetic cause of infant mortality, is a devastating neurodegenerative disease caused by reduced levels of Survival Motor Neuron (SMN) gene activity. Despite well-characterized aspects of the involvement of SMN in snRNP biogenesis, the gene circuitry affecting SMN activity remains obscure. Here, we use Drosophila as a model system to integrate results from large-scale genetic and proteomic studies, and bioinformatics analyses to define a unique SMN interactome to provide a basis for a better understanding of SMA. Such efforts not only help dissect the Smn biology but may also point to potential clinically relevant targets.

Reserved for Publication Footnotes

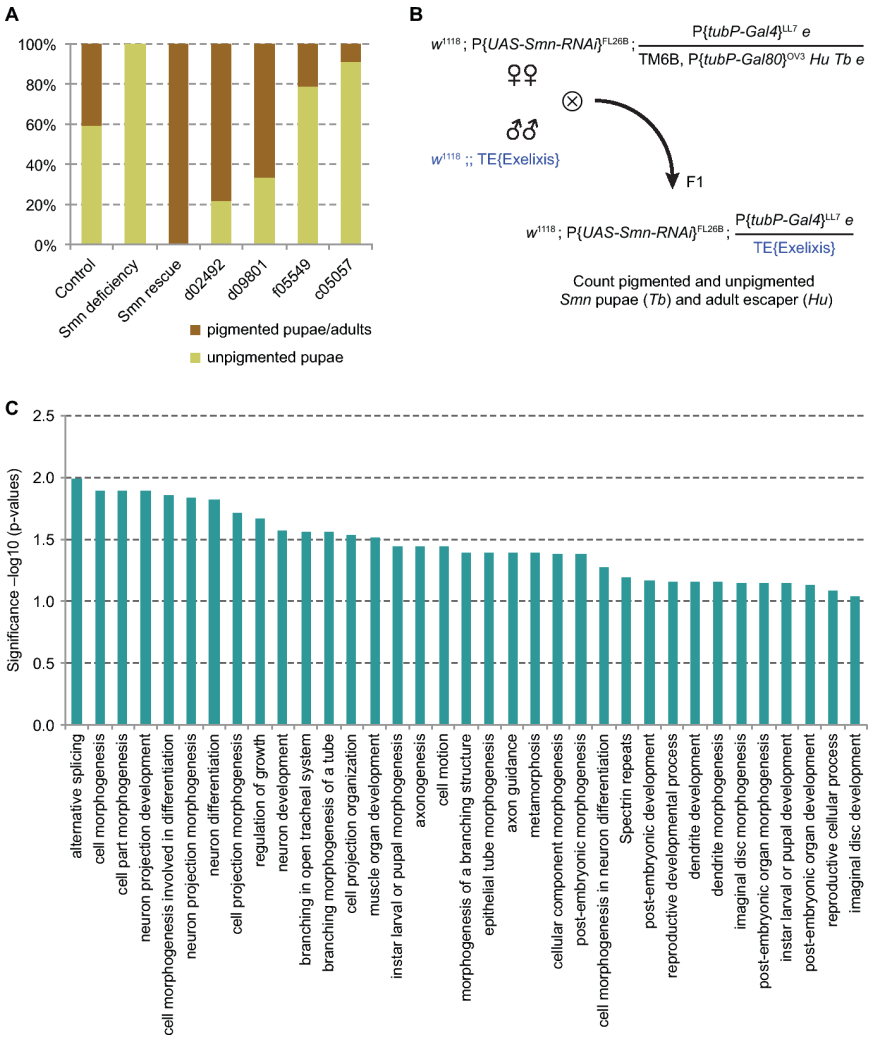


Fig. 1. Genetic modifiers of *Smn* using pupal lethality to screen the Exelixis collection of transposon insertions and their functional roles(A) *tubulinGAL4* (*tub-GAL4*) directed expression of an inducible *Smn*-RNAi construct (*UAS-Smn-RNAi*^{FL26B}) leads to a fully penetrant pupal lethality where approximately 40% of the pupae reach a pigmented developmental stage (Control). The remaining 60% die at an earlier unpigmented developmental stage. Introduction of an *Smn* deficiency into this background causes the entire population of pupae to die at the unpigmented stage (*Smn* deficiency), while ectopic *Smn* expression leads to survival to adulthood of the vast majority of pupae (*Smn* rescue). Introduction of previously isolated enhancers (d02492 and d09801) and suppressors (f05549 and c05057) of *Smn* (13) lead to quantitative changes in the fractions of pigmented vs. unpigmented pupae. (B) The screening strategy to identify genetic modifiers of the *Smn* pupal lethality phenotype using the Exelixis collection (illustrated for insertions on the 3rd chromosome). The lethal phase for all *Smn Tb*+ TE (16) pupae in individual test crosses are scored and compared to those observed in control crosses (more survival = enhancers, more lethality = suppressors). (C) *Drosophila* functional categories over-represented in the genetic modifier list. GO biological functions with the highest significance relate to known Smn functions such as alternative splicing or SMA affected processes (neuronal and muscular). Enrichment significance is expressed as the –log$_{10}$ (p-values).

intervention points that carry potential for so far unexplored clinical applications.

RESULTS

A genetic screen for modifiers of *Smn*-dependent lethality

We examined several *Smn*-RNAi strains to identify a hypomorphic *Smn* allele that could be used to model SMA in *Drosophila* more faithfully than alleles that completely abolish *Smn* function. We identified a transgenic strain, *UAS-Smn-RNAi*^{FL26B} (FL26B) that displays a less severe phenotype than the allele used in our previous screen (13). Specifically, expression of FL26B under the control of *tubulinGAL4* (*tubGAL4*::*FL26B*) results in late pupal lethality whereby approximately 50% of the pupae reach a more mature (pigmented) developmental stage prior to death (**Figure 1A**) than their less mature, unpigmented siblings.

We determined that this phenotype, measured by the ratio of pigmented to unpigmented pupae, is sensitive to *Smn* dosage, as reducing or increasing *Smn* copy number in the *tubGAL4*::*FL26B* genetic background resulted in enhancement or suppression, respectively (**Figure 1A**). In addition, the ability of wild type *Smn* (expressed by a *UAS-Smn-GFP* transgene) to rescue the lethality indicates that this phenotype does not result from off target RNAi effects. These results were corroborated using an independent *Smn* RNAi strain. Finally, we demonstrated that previously identified *Smn* modifiers altered the *Smn* RNAi phenotype in the expected fashion (**Figure 1A**). Together, these results demonstrate that the *tubGAL4*::*FL26B* phenotype is useful to detect changes in *Smn* functional activity and is thus suitable assay on which to base a modifier screen that will define and dissect the Smn genetic network.

Using this novel assay, we screened the Exelixis collection of genome-wide insertional mutations [https://drosophila.med.harvard.edu/ and (15, 16)] for dominant modification of the lethality associated with the *tubGAL4*::FL26B strain (see **Figure 1B** for scheme) and identified nearly 1600 candidate strains. To eliminate false positives, all interacting insertions were retested. Only those that were not lethal in *trans* with *tubGAL4* (13) and for which results could be clearly repeated were finally designated as modifiers. From this analysis, we identified 303 modifying strains (129 enhancers and 174 suppressors), which represents nearly 2% of the collection and a greater than tenfold increase in hit recovery in comparison to our previous screening method. As the genomic location of some insertions in the Exelixis collection may be near multiple coding regions, unambiguous gene assignments are not always possible. Given this consideration, we determined that these 303 insertions potentially affected 340 *Drosophila* genes. In most cases, single genes were affected by single insertions, though 14 genes scored in the screen were represented by 2 or more alleles. No gene could be assigned for 36 of the 303 insertions. Careful human to *Drosophila* homology mapping using a combination of

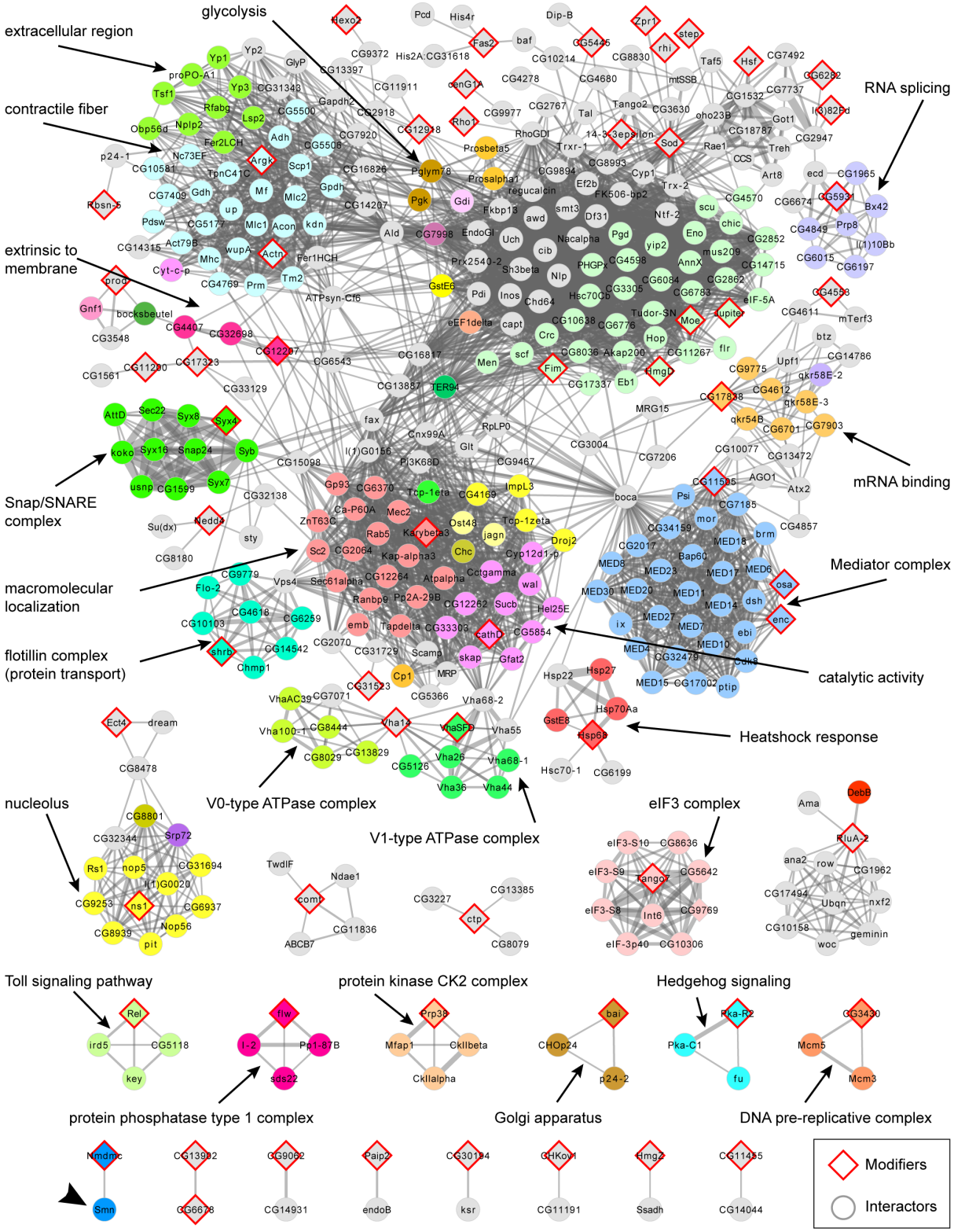


Fig. 2. The extended *Drosophila* genetic sub-network The sub-network of proteins connected to Smn and its genetic modifiers in the *Drosophila* Protein Interaction Map (DPiM). A total of 62 Smn genetic modifiers (diamonds with red border) are directly connected to 361 other proteins (circles), also known as first-degree neighbors through 3,800 interactions. The thickness of the gray lines connecting the proteins is proportional to the interaction score in the DPiM. Proteins belonging to individual clusters with GO term enrichment are shown with different colors. Proteins colored gray are part of clusters that are not enriched for any specific GO terms. Smn protein (indicated by an arrow head) itself is only connected to Nmdmc and shown as an interacting pair at the bottom.

several prediction algorithms show that out of the 340 *Drosophila* genes, at least 229 have human orthologs. Since a fraction of

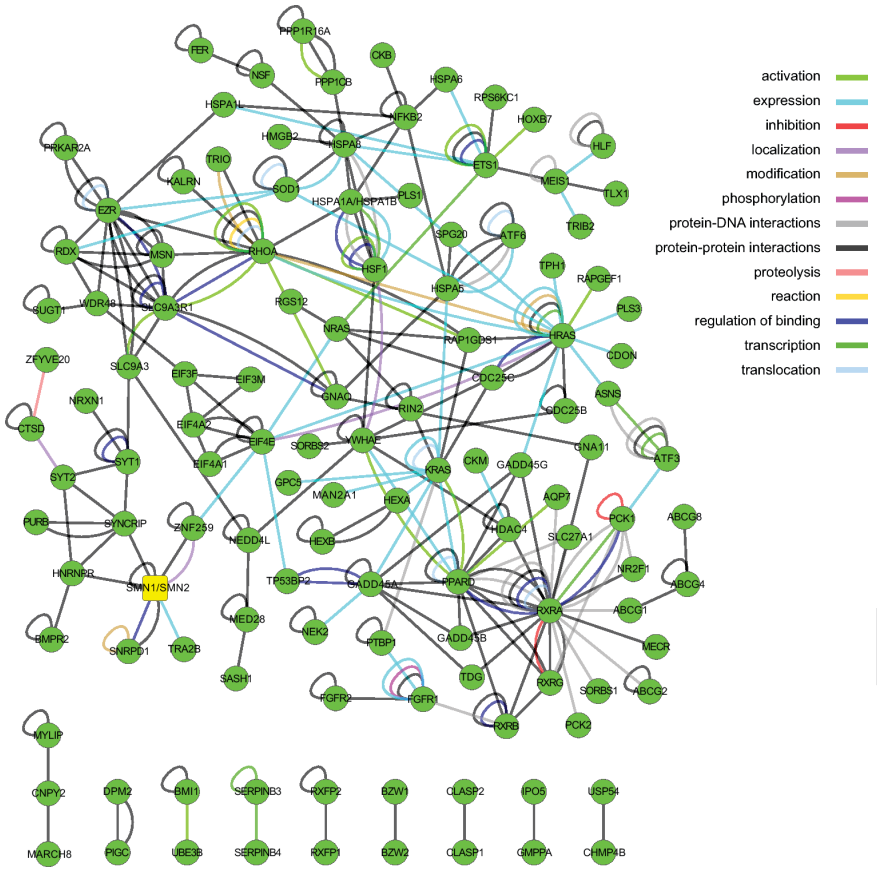


Fig. 3. Human Smn genetic modifiers network Ingenuity Pathway Analysis indicates that about one-third of the human orthologs of *Drosophila* Smn genetic modifiers (103 genes, green circles) are connected in a network involving 282 interactions with other modifiers and SMN1/SMN2 (yellow). Different types of interactions are indicated with distinct colored lines. A small number of modifiers (19 genes) have only interactions with other modifiers but not SMN1/SMN2 (4 pairs are two human orthologs of the same *Drosophila* gene). Nearly two-thirds of the modifiers (177 genes, not shown) have no interactions that would connect them to SMN1/SMN2 or other modifiers

these genes are represented by multiple paralogs, we identified a total of 322 human genes corresponding to the 229 *Drosophila* modifiers (see Materials and Methods, **Dataset S1**).

To assess the biological space covered by the novel *Smn* modifiers they were evaluated using the **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (DAVID) (17, 18) (http://david.abcc.ncifcrf.gov/), which identifies biological processes statistically overrepresented in the set of genetic modifiers. Analysis of the *Drosophila* SMN network with DAVID identified known SMN molecular activity (alternative splicing) and SMN-dependent processes ("neuron differentiation", "axon guidance", "axonogenesis", "muscle organ development" and "dendrite morphogenesis/development") (**Figure 1C**) (13, 19). The predominant processes enriched in the modifier set reflect broad effects on morphogenesis and development reinforcing the notion that Smn depletion has pleiotropic consequences.

Integration of *Drosophila* genetic and proteomic interactors

To determine whether the genetic modifiers are interconnected through physical interactions, we placed them in the context of the recently generated *Drosophila* Protein Interaction Map (DPiM) (14). We first retrieved the set of proteins that co-purify with *Drosophila* Smn in DPiM, and asked whether any of the modifiers belong to this Smn sub-network. From all *Drosophila* proteins tested, 8 form protein complexes with Smn and interestingly one, *NAD-dependent methylenetetrahydrofolate dehydrogenase* (*Nmdmc*), is also a genetic modifier. Expansion of the Smn subnetwork to include proteins that form complexes with each of the eight members in turn identified 35 additional proteins. Some of these proteomic interactors have been previously associated with Smn function (e.g. Gemin2, Gemin3 and several snRNPs), and nearly half (20/43) are known to be involved in mRNA processing, a functionality closely linked to the documented biochemical role of SMN (**Figure S1, Dataset S2**).

To better understand the biological functions identified by the Smn modifiers, we extracted the *Drosophila* protein complexes, which include the Smn genetic modifiers in DPiM. We found that 62 of the proteins corresponding to genetic modifiers passed the stringent statistical criteria necessary for inclusion in the DPiM interactome (**Materials and Methods**) which includes only the top 5% of the total interactions scored in the coAP/MS analysis. These 62 proteins were associated with 50 separate Markov clusters, a statistical definition of significantly associated proteins (20), each of which may define a functional protein complex (see **Figure 2** for the sub network of complexes identified by this analysis). Of these 50, we focused our attention on the 24 that are enriched for specific biological functions based on Gene Ontology (GO) terms – a system that provides a controlled vocabulary of terms for describing gene cellular and molecular functions (21). The majority of these clusters harbor a single modifier, but 9 contained two or more (**Table S3**).

Inspecting the GO biological function terms of these 24 clusters (annotated in Figure 2), we find that many contain annotated functions previously linked to *Smn* activity including RNA metabolism ("RNA splicing", "mRNA binding") [reviewed in (3)], translation control ("eIF3 complex") (22-25), endocytosis ("Snap/SNARE complex") (26, 27), and protein transport ("flotillin complex") (28-30). Importantly, several genetic modifiers fall within protein complexes whose functions have not been previously associated with *Smn* function. These include complexes with phosphatase and kinase activities as well as those involved in intracellular signaling ("Toll signaling pathway" and "Hedgehog signaling") (Figure 3B). Two independent loss-of-function alleles of the *Ect4/dsarm* gene suppress *Smn* dependent

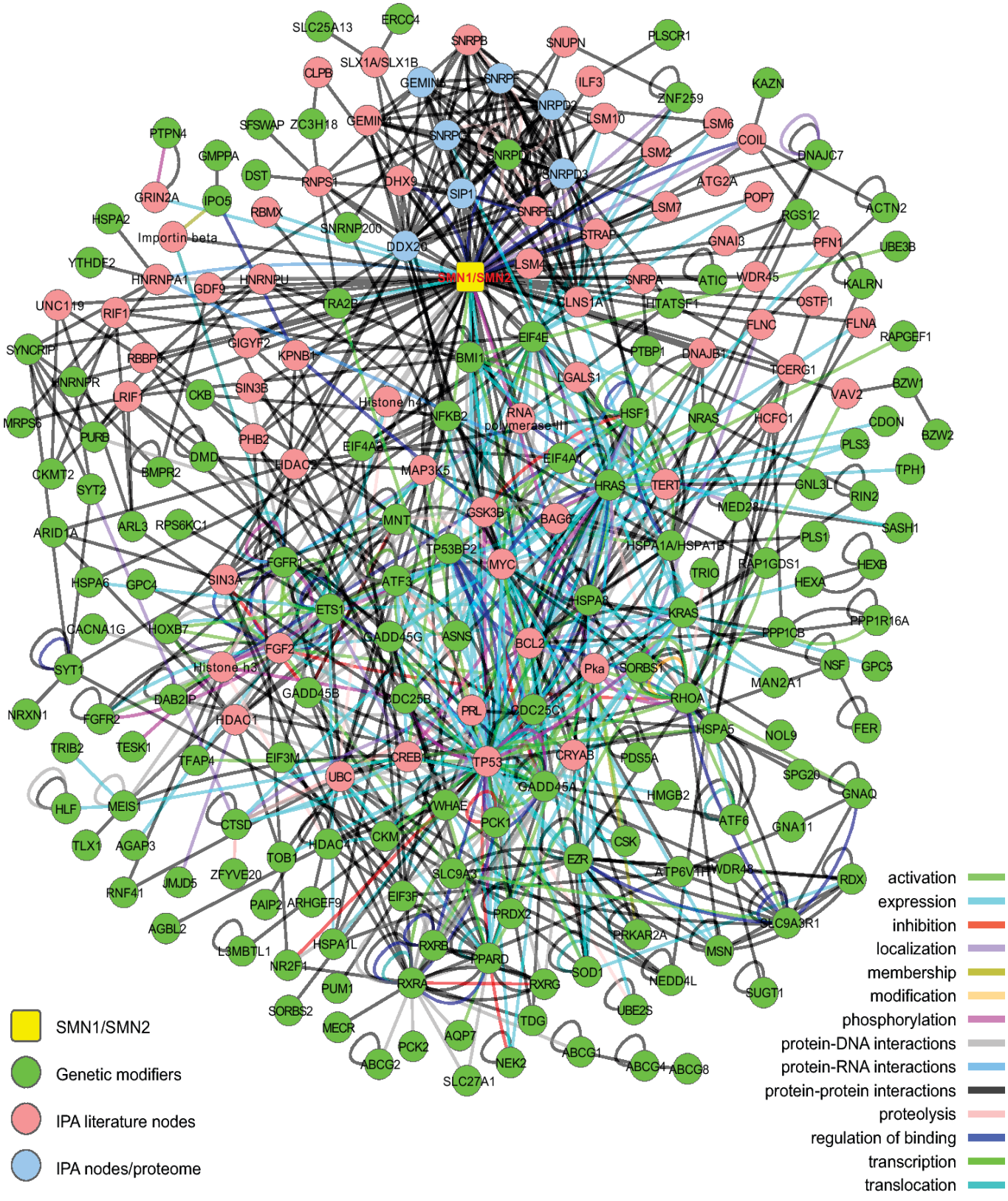


Fig. 4. Extended human Smn genetic sub-network Adding SMN1/2 first-degree neighbors to the network shown in Figure 3 generated an extended sub-network. In this interactome, 151 human orthologs of *Drosophila* Smn genetic modifiers (green circles) directly or indirectly connected to SMN1/SMN2. A total of 48 modifiers are connected to SMN1/SMN2 through 71 additional intermediate proteins (pink circles) from literature, seven among them (blue circles) were also identified in replicate SMN bait purifications in *Drosophila*. Different types of interactions are indicated with distinct colored lines

lethality. The recovery of *Ect4/dsarm* may provide additional evidence linking the Toll signaling pathway to *Smn* activity as it

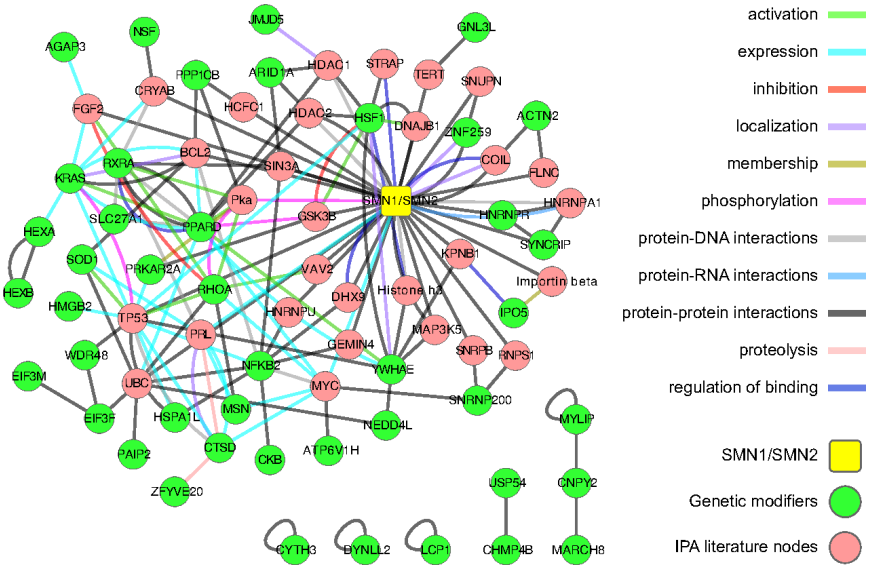
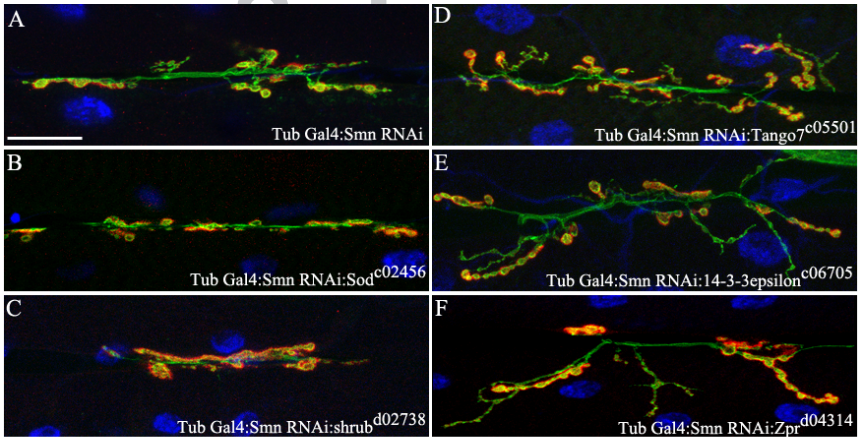


Fig. 5. Network of 36 high-priority genetic interactors of Smn The network shows 36 human orthologs of *Drosophila* Smn genetic modifiers (green circles) connected to SMN1/SMN2 in human. These 36 modifiers are present in *Drosophila* DPiM as well as human IPA based network and were selected for functional validation in NMJs. The intermediate proteins (pink circles) shown provide the shortest path to connect the modifiers to SMN1/SMN2. Different types of interactions are indicated with distinct colored lines.



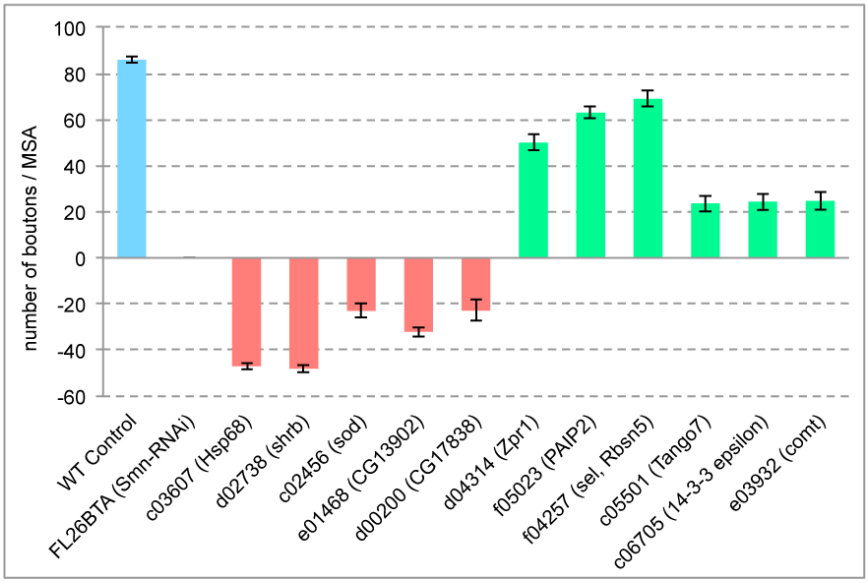


Fig. 6. Genetic modifiers of *Smn* regulate NMJ morphology. (A) An NMJ derived from muscle 6/7 of a *tubulinGAL4::UAS-Smn-RNAi*^{FL26B} (*tubGAL4::FL26B*) 3rd instar larva. (B, C) A reduction in NMJ size is observed upon introduction of enhancers c024569 (B) and d02738 (C) into the *tubGAL4::FL26B* background. (D, E, F) Introduction of a suppressors c05501, c06705 into this screening background leads to an increase in NMJ size, whereas suppressor d05711 (F) does not result in significant modification of the NMJ (G) Quantitation of bouton numbers/muscle in individuals of indicated genotypes, which include enhancers (red) and suppressors (green), normalized per muscle surface area (MSA) and expressed as percentage change as compared to Tub Gal4:Smn RNAi alone. The ANOVA multiple comparison test was used for statistical analysis of the bouton numbers/muscle. Significance P< 0.05. Scale Bar = 50 mm. n = 20. All preparations were stained with anti-HRP (red) and anti-Dlg (green). The muscle nucleus was labeled using DAPI.

encodes a Toll/interleukin-1 receptor homology (TIR) domain. Intriguingly, loss-of-function *Ect4/dsarm* mutations also suppress Wallerian degeneration phenotypes observed in *Drosophila* and mouse models (31). Together, these data suggest that the Wallerian degeneration pathway may also affect Smn pathobiology, an effect that may be mediated through Toll signaling. Hence this

approach both confirmed and expanded the functional categories and pathways associated with SMN.

To further explore the relationships of the 62 proteins and their functional context within DPiM, we carried out a first-degree neighbor analysis to identify other proteins directly connected in the network that may represent potential biochemical interactors. This retrieved 361 additional proteins that are linked to the 62 Smn modifier interactors (**Figure 2, Dataset S3**). These 361 proteins include 128 that are directly linked to at least two of the 62 modifiers (**Table S3**). A GO term (i.e. functionalities) analysis of these proteins reveals additional connections to the spliceosome, RNA binding and Snap/SNARE functions. Thus, considering genetic modifiers in the context of the DPiM provides us with a novel perspective of the diverse molecular functions that can modulate SMN activity *in vivo*.

Overlaying the genetic modifiers on the human interactome

To study the Smn genetic circuitry in the human context, we generated a human view of the genetic Smn interactome taking advantage of the manually curated source of human molecular interactions from IPA (Ingenuity® Systems, www.ingenuity.com). This database integrates human gene relationships derived from a variety of experimental approaches, including proteomic studies. Using the human Smn proteins and the 322 human genes corresponding to the genetic modifiers identified in *Drosophila* (see above and Materials and Methods), we used the IPA knowledge-base to derive a human SMN interaction network. Unlike DPiM, IPA is not limited to physical interactions thus allowing consideration of other functional interactions including, for example, expression, localization, modification, and regulation. Such an approach allowed us to evaluate potential indirect relationships between the modifiers and SMN, and uncover molecular functions beyond its canonical role in the SMN complex.

Based on the generated network, we found that orthologs of five modifiers HNRNPR, SNRPD1, SYNCRIP, TRA2B and ZNF259 are directly related to SMN1/2 (**Figure 3**). HNRNPR, SNRPD1 and SYNCRIP proteins physically interact with SMN1 and 2 and have a role in RNA splicing (32). Trab2 (also known as SFRS10 or Htra2-beta) was shown to regulate Smn2 protein levels by being a potent splicing enhancer (33). Finally, ZNF259 also known as ZPR1 was shown to be necessary for the localization of Smn1 to nuclear bodies (34) and more recently emerged as a key modifier of SMA pathology in patients (35). These findings support the relevance of the identified *Drosophila* modifiers in understanding the human pathways underlying SMA pathology.

Furthermore, 98 modifiers are indirectly related via these five interactors to human SMN. Together these 103 proteins representing one-third of the identified modifiers are interconnected in a human IPA database. In addition, we find another group of 19 proteins that make pair-wise functional interactions with other SMN genetic modifiers, but do not connect to the human Smn interactome. The remaining 177 proteins that are not connected in the human interactome (and the 19 that have pair-wise connections, 4 pairs of which are between two orthologs of the same *Drosophila* gene) potentially represent functions that have not been linked to SMN biology in human studies so far.

Expansion of the human SMN interactome beyond the 103 modifiers, by incorporating first-degree neighbor proteins of SMN in the database, connects an additional 48 modifiers (**Figure 4**). This expanded human SMN network contains intermediates that are known to associate physically with SMN (GEMINs, HNRNPs and LSM and SNRP family members) (3, 14) and signaling pathway elements known to affect SMN activity (FGF2, GSK3B, MAP3K5) (19, 36-39).

Validation of genetic modifiers at the larval NMJ

We chose to prioritize the modifiers for further functional characterization by using membership within both the Smn modifier network in DPiM (**Figure S1**) and the expanded IPA Human SMN network (**Figure 4**) as the primary criterion. A total of 36 genes are shared between these interactomes (**Figure 5, Dataset S4**). The list includes 4 previously analyzed modifiers (*Actn*, *Moes*, *Fim*, *cut up*) (13, 27), 13 enhancers, (*Sod*, *Hsp68*, *Hsf*, *step*, *CG17838*, *ns1*, *shrb*, *VhaSFD*, *Rel*, *Hexo2*, *osa*, *CG13902*, *cathD*) and 19 suppressors (*CG30194*, *Nedd4*, *Pka-R2*, *Rho1*, *Tango7*, *Argk*, *14-3-3-epsilon*, *Zpr1*, *CG9769*, *cenG1A*, *flw*, *comt*, *CG9062*, *l(3)72Ab*, *Karybeta3*, *HmgZ*, *Rbsn-5*, *sel*, *Paip2*).

Our previous analyses (13) indicated a strong correlation between the strength of the lethal *Smn* phenotype with the severity of NMJ abnormalities. Therefore, examination of the effects of *Drosophila* modifiers on the *Smn* NMJ phenotype was used to validate their role in *Drosophila* and prioritize the corresponding orthologs for further investigation in vertebrate model systems. We used NMJ assays (13, 19) to sample the ability of a subset of these modifiers to alter the *tubGAL4*::FL26B NMJ phenotype. Examination of third instar larvae carrying a combination of *tubGAL4*::FL26B and each of 20 modifier strains revealed that 11 out of the 20 (55%) strains reveal a statistically significant change in the number of synaptic boutons and are modifiers of the *Smn* NMJ phenotype (**Figure 6**).

Effect of genetic modifiers on *Smn* protein levels and localization

Given that the severity of the disease phenotypes, in both patients and *Drosophila* models, correlates with SMN proteins levels, we examined whether the prioritized genes affected SMN levels in *Drosophila* S2R+ cells (40), the same cell line used to generate DPiM. We used an image-based analysis (37) to quantify SMN protein levels in S2R+ cells expressing inducible FLAG-HA tagged constructs corresponding to 21 *Smn* modifying genes available from the Universal Proteomics Resources (41). Untransfected cells within the same wells were used as controls. Surprisingly, we found that none of these ectopically expressed modifier genes significantly altered total *Smn* protein expression (**Figure S2A**). Since Smn is localized in both the cytoplasm and the nucleus, we also used this assay to evaluate whether any of these modifiers altered its distribution between these two compartments. We found seven modifiers significantly increased the nuclear Smn levels (**Figure S2B** and **Dataset S5**), consistent with the notion that some modifiers from the screen, which affect Smn lethality and NMJ phenotype, may directly affect Smn distribution between the nucleus and cytoplasm. It is worth noting that, a recent study (42) showed that mutant superoxide dismutase-1 (SOD1), known to cause familial ALS, alters the sub-cellular localization of the SMN protein and disrupts its recruitment to Cajal bodies thereby preventing the formation of nuclear 'gems'. *Sod* was identified in our screen as an enhancer and was also shown to affect NMJ phenotype (**Figure 6**). A subset of modifiers does not alter either Smn levels or its localization. How these modifiers affect the functional Smn remains to be explored. Given these results, however, it is important to note that small changes in SMN function may have an important biological impact given that the severity of clinical manifestation in SMA patients correlates with small changes in SMN expression (1).

DISCUSSION

Different animal models for SMA-associated neuromuscular defects contributed significantly to a better understanding of the Spinal Muscular Atrophy etiology and genetics over the last few years. However, despite the well-characterized role for SMN in snRNP biogenesis, the links between its molecular function and the defects observed in SMA patients remain unclear. One of the key features of SMA is that the severity of the disease is dependent on SMN dosage, prompting the development of therapeutic strategies designed to increase SMN protein levels in patients. Still, it is essential to identify alternative approaches to modulate SMN activity. For this purpose, genetically tractable invertebrate

systems may help to identify so far undiscovered elements of the SMN genetic circuitry. In particular, these organisms provide more flexible avenues to investigate the poorly understood role of SMN at the NMJ.

We have used *Drosophila* as our experimental system and previously described a genetic screen which uncovered a small number of *Smn* modifiers (13) of a strong loss of function mutant phenotype. In this screen, we identified functional links between *Smn* and the FGF pathway (13, 19), a relationship corroborated and extended by recent evidence in a severe mouse model of SMA, which demonstrated widespread alterations of the FGF-system in both muscle and spinal cord (38).

The relatively small number of modifiers recovered suggested a more sensitive genetic screen could provide extended information about the *Smn* genetic network. Our assessment of the lethal phase exhibited by a mild loss of function *Smn* RNAi allele, which more closely resembles the SMA hypomorphic condition, provided us with a more sensitive and quantifiable assay for genetic interaction. In comparison to our previous results, the RNAi-based screen described here provided us with a broader spectrum of modifiers including those related to the canonical role of *Smn* in snRNP biogenesis as well as additional elements of FGF and BMP signaling (13). Our careful mining of the screening modifier list based on functional term enrichment, and interactome analysis both in *Drosophila* and human, suggest that loss of *Smn* function may impact a range of developmental and maintenance-related programs of the whole neuromuscular system, including synaptic vesicle recycling, ion channels and signaling pathways that regulate intrinsic cellular functions. Finally, this analysis also uncovered biological processes not previously associated with *Smn*.

Among the newly recovered genes, many are associated with RNA metabolism; however, the majority is not involved with canonical SMN activity of snRNP biogenesis and includes factors involved in transcription, post-transcriptional modifications, RNA transport and translation regulation. Intriguingly, *CG17838* is the *Drosophila* homolog of two closely related vertebrate RNA-binding proteins, hnRNP-R and SYNCRIP/hnRNP-Q, both of which bind to SMN in a yeast two-hybrid assay (32) and localize to mRNA containing granules that are transported in cultured neurons (28, 32, 43). Since both SMN and hnRNP-R affect localization of mRNA in axons (44, 45), this could have profound consequences on local translation in neurons (45).

Given the complexity of motor neuronal sub-cellular domains and their distance from the neuromuscular synapses, local regulation of the translation of synaptic proteins is likely to be important in synaptic plasticity and neurological diseases. In fact, many mRNA binding proteins (RBPs) that function as key regulators of local RNA translation are associated with neurological diseases, including FMRP in Fragile X Syndrome, ATXN-2 in Spinocerebellar Ataxia, and TDP-43, FUS (fused in sarcoma), ANG and ATXN-2 in Amyotrophic Lateral Sclerosis. Consistent with a possible role for *Smn* in affecting local translation, we recovered *pumilio* and *eIF-4E*, which are thought to be a part of the local translational apparatus in neuromuscular synapses (46). Furthermore, we recovered another translation regulator, *eIF-4A*, which negatively regulates BMP signaling components. Components of BMP signaling pathway have been shown to play a role in retrograde signaling in the NMJ (47, 48). Our results support the relationship between Smn and local translation and also provide an additional link to the retrograde signaling present in the neuromuscular system. Interestingly, perturbation of RNA translational control may result in defects in endocytosis (49, 50), a process that has been suggested play a key role in neurodegenerative diseases, including Alzheimer's (51) and Huntington's diseases (52). Consistent with this notion, aberrant synaptic vesicle release at the NMJs in severe SMA mice may be evidence of impaired synaptic vesicle dynamics and/or abnormal active zone architecture (53). Further supporting a link between endocytosis and Smn (27), we identified *Synaptotagmin1, Synaptotagmin-alpha*, *Syntaxin4* and *comatose,* the *Drosophila* homolog of N-ethylmaleimide sensitive factor (NSF), which are core components of synaptic vesicle recycling. We also recovered genes that are directly required for synaptic transmitter release, such as *methuselah*, or indirectly, such as *bruchpilot*, which plays a role in constructing the active zone.

Though many of the recovered genes broadly impact the neuromuscular system, a subset includes the *Drosophila* homologs for several disease related genes, including *Nrx-1* (schizophrenia and Autism Spectrum Disorders) (54, 55), *Dystrophin* (Duchenne's Muscular Dystrophy) (56, 57), *Superoxide dismutase* (42, 58), *RhoA*, (Amyotrophic Lateral Sclerosis) (59) and *Ect4*/*dSarm* (Wallerian degeneration) (31). Our recovery of these genes suggests the genetic network identified by our screen may overlap, perhaps significantly, the genetic networks impacted by other human neurological disorders. If true, the use of *Drosophila* to explore other neurological disease networks via genetic screens, combined with the integration of additional genome-wide approaches, could identify common therapeutic targets which could potentially be tested in other disease models.

Since such genetic modifier screens are very sensitive and are able to recover a large number of modifiers that span a broad range of molecular functions, prioritization of candidates for further validation is essential. Here, bioinformatics mining allowed us to assemble a list of 36 *Drosophila* genes with human homologs for continued investigation. The majority of these tested genes showed a functional role in the structure and/or development of the NMJ in *Drosophila*, and some can alter the distribution of Smn in S2R+ cells, making them good candidates to pursue in vertebrate models of SMA. In addition, candidates may be drawn from a pool of modifiers that include members of signaling pathways such as GPCR, Kinases and Proteases, which are considered to be plausible small molecule targets, or secreted or membrane proteins, which may be targeted by antibodies. Our results thus provide an extensive list of novel genes and pathways that have now been functionally linked to an *Smn*-dependent phenotype and therefore represent potentially novel therapeutic targets.

MATERIALS AND METHODS

Drosophila stocks and culture

All *Drosophila* stocks were maintained on standard *Drosophila* medium at 25°C. The generation of the *Smn* alleles and constructs used in this study (*Smn*^{X7}, *UAS-Smn-RNAi*^{FL26B}, *UAS-Smn-GFP*) were originally described in (13). The tubulinGAL4 and TM6B, Tb Hu tubulinGAL80 chromosomes used to generate the screening stock were obtained from the Bloomington *Drosophila* Stock Center (Bloomington, IN). The Exelixis Collection is housed in the Artavanis-Tsakonas laboratory in the Department of Cell Biology at Harvard Medical School (Boston, MA).

Genetic Modifier Screen

Individual strains from the Exelixis Collection were tested for the ability to genetically modifiy the *tubGAL4*-induced *UAS-Smn-RNAi*^{FL26B} pupal lethal phenotype by mating 3-5 males of the strain to 3 females of the *w*; *UAS-Smn-RNAi*^{FL26B}; *tubGAL4*/TM6B, *Hu Tb tubGAL80* screening stock. After two days, adults were transferred to a fresh vial to create a duplicate cross and to maintain optimal culture density. Accordingly, adults were discarded from the duplicate after an additional two days had passed. 15 days after being initiated, crosses were scored by counting the number of pigmented and unpigmented Tb+ pupae along with any Hu+ adult escapers. The ratio of unpigmented pupae to the total of pigmented pupae + escapers adults was compared to that derived from control crosses using males from the isogenic strain in which the Exelixis Collection was generated. Two control crosses were performed for each set of approximately 100 strains that were tested. Crosses that failed to produce 40 experimental animals were repeated as above. A change in the ratio of unpigmented to pigmented individuals of 20% corresponded to an approximate 1.5 standard deviation from the mean. Enhancers were defined as those mutations causing a reduction in this percentage (≤30%), while those that increased this percentage (≥70%) were classified as suppressors.

Gene Assignments for the Exelixis Collection of Transposon Insertions

Data for *Drosophila* genes and Exelixis transposon insertion sites were obtained from FlyBase version 5.39, which was current as of August 2011. Of the 15,952 Exelixis stocks screened, 14,621 stocks were mapped in FlyBase to 15,326 transposons with specific insertion sites within the *Drosophila* genome. To determine the coordinates of insertion sites of transposons present in the remaining 331 strains, sequences from the region flanking the insertion sites (15) were searched against the *Drosophila* genome using the blastn program of the Basic Local Alignment Search Tool [BLAST, (60)]. The insertion sites of the transposons were then used to create gene assignments according to the following criteria: a transposon was considered to map to a particular gene if its insertion site is located in the transcription unit of the gene itself or within either 1 kb upstream of the transcription start site or 100 bp downstream of the transcription termination site.

Mapping *Drosophila* genes to human orthologs

FlyBase version v5.39 identifies 15,233 *Drosophila* genes, which were iteratively mapped to human orthologs using predictions made by several prediction algorithms. Multiple predictions were combined into a single prediction by ordering the algorithms based upon lowest false positive and highest false negative rates [see (61)], and choosing the first prediction. The methods used (in order) were inParanoid version 7 (62), orthoMCL version 5 (63), Homologene build 65 (64) and, orthoMCL version 2 (65). The inParanoid predictions were selected using a probability score of 0.4. As a result, of the 15,233 *Drosophila* genes considered, 6,821 could be mapped to 6,703 human gene ids. This dataset was used to assign the human orthologs of *Drosophila* Smn modifiers as shown in Dataset S1.

Function and network analysis

The functional enrichment Gene Ontology terms of the *Drosophila* genetic modifiers was assessed using EASE statistics available through the Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics Resources using the Exelixis collection as a reference set (17, 18). Human and *Drosophila*, protein-protein and genetic interactions were visualized and analyzed (first neighbors) using Cytoscape (66). Cytoscape BinGo plugin was used to evaluate the Functional categories of the retrieved clusters in the *Drosophila* sub-networks. The human network was generated through the use of IPA (Ingenuity Systems, www.ingenuity.com) and further visualized and mined in Cytoscape.

Neuromuscular junction analyses

Third instar larvae were dissected in cold 1X phosphate buffered saline (PBS) and fixed at room temperature (RT) for 20 min. in 4% Paraformaldehyde (PFA). The samples were washed in 0.1% Triton-X100 in PBS (PTX) and incubated overnight at 4°C with primary antibody. The primary antibody was washed off with PTX at RT. The samples were incubated at RT with secondary antibody for 90 min. This was followed by PTX wash, and the tissues were mounted in Vectashield Mounting Media with DAPI (Vector Laboratories). Bouton numbers were counted using a Nikon TE2000 microscope, based on the Discs large and anti-HRP staining in the A3 segment muscle 4 as indicated. The muscle area for every animal was measured, and no significant difference was observed among different genotypes. At least 20-25 animals of each genotype were dissected for the bouton analysis. The ANOVA multiple comparison test was used for statistical analysis of the bouton number/muscle.

Microscopy

All images were collected with a Nikon C1si spectral point scanning confocal connected to a Nikon TE2000 inverted microscope equipped with DIC, phase, and epi-fluorescence optics, 40x Plan Fluor NA 1.4 objective lens and the Perfect Focus System for continuous maintenance of focus. 100mW mercury arc lamp illumination for viewing fluorescence by eye, and confocal scanning using Melles Griot solid state diode lasers: 405nm, 488nm (10mW), and 561nm (10mW). The image acquisition software used was Nikon EZ-C1. All samples were mounted and imaged in Vectashield mouting medium with DAPI (Vector Laboratories) at room temperature. Adobe Photoshop CS5 was used to pseudocolor images.

Analysis of Smn levels in S2R+ cells

Drosophila S2R+ cells (40), a derivative of Schneider S2 cells, were cultured in Schneiders *Drosophila* medium (Gibco) with 10% fetal bovine serum, 100 U/ml of penicillin, and 100 µg/ml of streptomycin at 25°C. FLAG-HA constructs (Guruharsha et al., 2011) were transfected using TransIT-2020 (Mirus). One day after transfection, plasmid expression was induced with 0.35 mM $CuSO_4$ overnight. Harvested cells were plated on concanavalin A (0.5 mg/ml, Sigma) -coated plate and fixed at room temperature for 20 min in 4% paraformaldehyde (Electron microscopy sciences). The cells were washed in PBS-DT (0.3% sodium deoxycholate, 0.3% Triton X-100 in PBS) and incubated overnight at 4°C with rabbit anti-Smn (1:2000) (Sen et al., 2011) and mouse anti-Flag (1:1000, Sigma). After washes in PBS-DT, the cells were incubated with mouse Alexa 488- and rabbit Alexa 568-conjugated secondary antibodies (1:500, Molecular Probes), followed by washing in PBS-T (0.1% Triton X-100 in PBS). The samples were mounted in VECTASHIELD mounting medium with DAPI (Vector Laboratories).

ACKNOWEDGEMENTS

This work was initially supported by a grant from the SMA Foundation to S. Artavanis-Tsakonas, A.C. Hart, and D. van Vactor. Additional support from Spinal Muscular Atrophy Program Project (P01N5066888) to S. Artavanis-Tsakonas, L. Rubin, A.C. Hart, and D. van Vactor. Work on DPiM project was supported by a grant from the National Institutes of Health (NIH 5RO1HG003616) to S. Artavanis-Tsakonas. A. Sen was supported by a postdoctoral fellowship from the Families of Spinal Muscular Atrophy. The authors thank Nina Makhortova for help with cell imaging and the Nikon Imaging Center at Harvard Medical School for help with light microscopy.

1. Lefebvre S, *et al.* (1997) Correlation between severity and SMN protein level in spinal muscular atrophy. *Nat Genet* 16(3):265-269.
2. Workman E, Kolb SJ, & Battle DJ (2012) Spliceosomal small nuclear ribonucleoprotein biogenesis defects and motor neuron selectivity in spinal muscular atrophy. *Brain Res* 1462:93-99.
3. Burghes AH & Beattie CE (2009) Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci* 10(8):597-609.
4. Lefebvre S, *et al.* (1995) Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* 80(1):155-165.
5. Lorson CL, Hahnen E, Androphy EJ, & Wirth B (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc Natl Acad Sci U S A* 96(11):6307-6311.
6. Monani UR, *et al.* (1999) A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet* 8(7):1177-1183.
7. Feldkotter M, Schwarzer V, Wirth R, Wienker TF, & Wirth B (2002) Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet* 70(2):358-368.
8. Schrank B, *et al.* (1997) Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc Natl Acad Sci U S A* 94(18):9920-9925.
9. Chan YB, *et al.* (2003) Neuromuscular defects in a Drosophila survival motor neuron gene mutant. *Human molecular genetics* 12(12):1367-1376.
10. Chan YB, *et al.* (2003) Neuromuscular defects in a Drosophila survival motor neuron gene mutant. *Hum Mol Genet* 12(12):1367-1376.
11. Miguel-Aliaga I, Chan YB, Davies KE, & van den Heuvel M (2000) Disruption of SMN function by ectopic expression of the human SMN gene in Drosophila. *FEBS Lett* 486(2):99-102.
12. Rajendra TK, *et al.* (2007) A Drosophila melanogaster model of spinal muscular atrophy reveals a function for SMN in striated muscle. *J Cell Biol* 176(6):831-841.
13. Chang HC, *et al.* (2008) Modeling spinal muscular atrophy in Drosophila. *PLoS ONE* 3(9):e3209.
14. Guruharsha KG, *et al.* (2011) A Protein Complex Network of Drosophila melanogaster. *Cell* 147(3):690-703.
15. Thibault ST, *et al.* (2004) A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac. *Nat Genet* 36(3):283-287.
16. Artavanis-Tsakonas S (2004) Accessing the Exelixis collection. *Nat Genet* 36(3):207.
17. Huang da W, Sherman BT, & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44-57.
18. Huang da W, Sherman BT, & Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37(1):1-13.
19. Sen A, *et al.* (2011) Modeling spinal muscular atrophy in Drosophila links Smn to FGF signaling. *J Cell Biol* 192(3):481-495.
20. Enright AJ, Van Dongen S, & Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30(7):1575-1584.
21. Ashburner M, *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25(1):25-29.
22. Sharma A, *et al.* (2005) A role for complexes of survival of motor neurons (SMN) protein with gemins and profilin in neurite-like cytoplasmic extensions of cultured nerve cells. *Experimental cell research* 309(1):185-197.
23. Leung KM, *et al.* (2006) Asymmetrical beta-actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nat Neurosci* 9(10):1247-1256.
24. Al-Ramahi I, *et al.* (2007) dAtaxin-2 mediates expanded Ataxin-1-induced neurodegeneration in a Drosophila model of SCA1. *PLoS Genet* 3(12):e234.
25. Piazzon N, *et al.* (2008) In vitro and in cellulo evidences for association of the survival of motor neuron complex with the fragile X mental retardation protein. *The Journal of biological chemistry* 283(9):5598-5610.
26. Oprea GE, *et al.* (2008) Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. *Science* 320(5875):524-527.
27. Dimitriadi M, *et al.* (2010) Conserved genes act as modifiers of invertebrate SMN loss of function defects. *PLoS Genet* 6(10):e1001172.
28. Rossoll W, *et al.* (2003) Smn, the spinal muscular atrophy-determining gene product, modulates axon growth and localization of beta-actin mRNA in growth cones of motoneurons. *The Journal of cell biology* 163(4):801-812.
29. Zhang HL, *et al.* (2003) Active transport of the survival motor neuron protein and the role of exon-7 in cytoplasmic localization. *J Neurosci* 23(16):6627-6637.
30. Bowerman M, Beauvais A, Anderson CL, & Kothary R (2010) Rho-kinase inactivation prolongs survival of an intermediate SMA mouse model. *Human molecular genetics* 19(8):1468-1478.
31. Osterloh JM, *et al.* (2012) dSarm/Sarm1 is required for activation of an injury-induced axon death pathway. *Science* 337(6093):481-484.
32. Rossoll W, *et al.* (2002) Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Human molecular genetics* 11(1):93-105.
33. Hofmann Y, Lorson CL, Stamm S, Androphy EJ, & Wirth B (2000) Htra2-beta 1 stimulates an exonic splicing enhancer and can restore full-length SMN expression to survival motor

neuron 2 (SMN2). *Proc Natl Acad Sci U S A* 97(17):9618-9623.
34. Gangwani L, Mikrut M, Theroux S, Sharma M, & Davis RJ (2001) Spinal muscular atrophy disrupts the interaction of ZPR1 with the SMN protein. *Nature cell biology* 3(4):376-383.
35. Ahmad S, Wang Y, Shaik GM, Burghes AH, & Gangwani L (2012) The zinc finger protein ZPR1 is a potential modifier of spinal muscular atrophy. *Human molecular genetics* 21(12):2745-2758.
36. Claus P, Bruns AF, & Grothe C (2004) Fibroblast growth factor-2(23) binds directly to the survival of motoneuron protein and is associated with small nuclear RNAs. *Biochem J* 384(Pt 3):559-565.
37. Makhortova NR, *et al.* (2011) A screen for regulators of survival of motor neuron protein levels. *Nat Chem Biol* 7(8):544-552.
38. Hensel N, *et al.* (2012) Analysis of the fibroblast growth factor system reveals alterations in a mouse model of spinal muscular atrophy. *PLoS ONE* 7(2):e31202.
39. Kwon JE, Kim EK, & Choi EJ (2011) Stabilization of the survival motor neuron protein by ASK1. *FEBS Lett* 585(9):1287-1292.
40. Yanagawa S, Lee JS, & Ishimoto A (1998) Identification and characterization of a novel line of Drosophila Schneider S2 cells that respond to wingless signaling. *J Biol Chem* 273(48):32353-32359.
41. Yu C, *et al.* (2011) Development of expression-ready constructs for generation of proteomic libraries. *Methods Mol Biol* 723:257-272.
42. Kariya S, *et al.* (2012) Mutant superoxide dismutase 1 (SOD1), a cause of amyotrophic lateral sclerosis, disrupts the recruitment of SMN, the spinal muscular atrophy protein to nuclear Cajal bodies. *Human molecular genetics* 21(15):3421-3434.
43. Gu W, Pan F, Zhang H, Bassell GJ, & Singer RH (2002) A predominantly nuclear protein affecting cytoplasmic localization of beta-actin mRNA in fibroblasts and neurons. *The Journal of cell biology* 156(1):41-51.
44. Glinka M, *et al.* (2010) The heterogeneous nuclear ribonucleoprotein-R is necessary for axonal beta-actin mRNA translocation in spinal motor neurons. *Human molecular genetics* 19(10):1951-1966.
45. Fallini C, *et al.* (2011) The survival of motor neuron (SMN) protein interacts with the mRNA-binding protein HuD and regulates localization of poly(A) mRNA in primary motor neuron axons. *J Neurosci* 31(10):3914-3925.
46. Menon KP, *et al.* (2004) The translational repressor Pumilio regulates presynaptic morphology and controls postsynaptic accumulation of translation factor eIF-4E. *Neuron* 44(4):663-676.
47. Aberle H, *et al.* (2002) wishful thinking encodes a BMP type II receptor that regulates synaptic growth in Drosophila. *Neuron* 33(4):545-558.
48. McCabe BD, *et al.* (2003) The BMP homolog Gbb provides a retrograde signal that regulates synaptic growth at the Drosophila neuromuscular junction. *Neuron* 39(2):241-254.
49. Gibbings DJ, Ciaudo C, Erhardt M, & Voinnet O (2009) Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. *Nature cell biology* 11(9):1143-1149.
50. Lee YS, *et al.* (2009) Silencing by small RNAs is linked to endosomal trafficking. *Nature cell biology* 11(9):1150-1156.
51. Wu F & Yao PJ (2009) Clathrin-mediated endocytosis and Alzheimer's disease: an update. *Ageing Res Rev* 8(3):147-149.
52. Trushina E, *et al.* (2006) Mutant huntingtin inhibits clathrin-independent endocytosis and causes accumulation of cholesterol in vitro and in vivo. *Human molecular genetics* 15(24):3578-3591.
53. Kong L, *et al.* (2009) Impaired synaptic vesicle release and immaturity of neuromuscular junctions in spinal muscular atrophy mice. *J Neurosci* 29(3):842-851.
54. Gauthier J, *et al.* (2011) Truncating mutations in NRXN2 and NRXN1 in autism spectrum disorders and schizophrenia. *Hum Genet* 130(4):563-573.
55. Rujescu D, *et al.* (2009) Disruption of the neurexin 1 gene is associated with schizophrenia. *Human molecular genetics* 18(5):988-996.
56. Koenig M, *et al.* (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organization of the DMD gene in normal and affected individuals. *Cell* 50(3):509-517.
57. Monaco AP, *et al.* (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* 323(6089):646-650.
58. Rosen DR, *et al.* (1993) Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362(6415):59-62.
59. Lesnick TG, *et al.* (2008) Beyond Parkinson disease: amyotrophic lateral sclerosis and the axon guidance pathway. *PLoS ONE* 3(1):e1449.
60. Altschul SF, *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* 25(17):3389-3402.
61. Chen F, Mackey AJ, Vermunt JK, & Roos DS (2007) Assessing performance of orthology detection strategies applied to eukaryotic genomes. *PLoS ONE* 2(4):e383.
62. Ostlund G, *et al.* (2010) InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res* 38(Database issue):D196-203.
63. Chen F, Mackey AJ, Stoeckert CJ, Jr., & Roos DS (2006) OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic acids research* 34(Database issue):D363-368.
64. Sayers EW, *et al.* (2012) Database resources of the National Center for Biotechnology Information. *Nucleic acids research* 40(Database issue):D13-25.
65. Li L, Stoeckert CJ, Jr., & Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* 13(9):2178-2189.
66. Shannon P, *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498-2504.